On the **Empirical Evaluation** of **Fault Localization** Techniques for **Spreadsheets**

Birgit Hofer, Franz Wotawa, Elisabeth Getzner

TU Graz

André Riboira, Rui Abreu

U.PORTO

16th Int. Conference on Fundamental Approaches to Software Engineering (FASE), March 2013, Rome, Italy

Spreadsheet Debugging

- Spreadsheet users outnumber programmers
- Basis for decisions
- Error prune
- No engineering discipline
- Hard task to debug spreadsheets

→ Use software debugging techniques

Outline

- Running Example & Definitions
- Debugging Methods
  - Spectrum-Based Fault Localization
  - Spectrum-Enhanced Dynamic Slicing
  - Constraint-Based Debugging
- Evaluation
- Future Work & Conclusion

Running Example

Faulty Spreadsheet

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	23	15	$345,00
3	Jones	35	34	69	17	$1.173,00
4	**Total**	58	65	92		

Formula View

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	=SUM(B2)	15	=D2*E2
3	Jones	35	34	=SUM(B3:C3)	17	=D3*E3
4	**Total**	=SUM(B2:B3)	=SUM(C2:C3)	=SUM(D2:D3)		

Basic definitions

- Spreadsheet language similar to Microsoft Excel
- Input cells
- Output cells
- Intermediate cells
- Test case
  - I = {B2=23, C2=31, E2=15, B3=35, C3=34, E3=17}
  - O = {D4=123, F2=810, F3=1173}

	A	B	C	D	E	F	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**	**Gross Pay**
2	Green	23	31	23	15	$345,00	=D2*E2
3	Jones	35	34	69	17	$1.173,00	=D3*E3
4	**Total**	58	65	92			

Program debugging: execution traces, slices
This work: cones

$$\mathrm{CONE}(c) = c \cup \bigcup_{c' \in \rho(c)} \mathrm{CONE}(c')$$

- The function *ρ(c)* returns all cells referenced in *c*.

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	=SUM(B2)	15	=D2*E2
3	Jones	35	34	=SUM(B3:C3)	17	=D3*E3
4	**Total**	=SUM(B2:B3)	=SUM(C2:C3)	=SUM(D2:D3)		

- CONE(F2) = {B2, D2, E2, F2}
- CONE(D4) = {B2, D2, B3, C3, D3, D4}

Faults where ∩ of cones does not work

- Several faults

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	=SUM(B2:C2)	15	=D2*C2
3	Jones	35	34	=SUM(B3:C3)	17	=D3*C3
4	**Total**	=SUM(B2:B3)	=SUM(C2:C3)	=SUM(D2:D3)		

  - Cone(F2) = {B2, C2, D2, F2}
  - Cone(F3) = {B3, C3, D3, F3}

- Single wrong output cell

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	=SUM(B2:C2)	15	=D2*E2
3	Jones	35	34	=SUM(B3:C3)	17	=D3*E3
4	**Total**	=SUM(B2:B3)	=SUM(C2:C3)	=SUM(D3:D3)		

  - Cone(F3) = {B3, C3, D3, D4, F3}

Outline

- Running Example & Definitions
- Debugging Methods
  - Spectrum-Based Fault Localization
  - Spectrum-Enhanced Dynamic Slicing
  - Constraint-Based Debugging
- Evaluation
- Future Work & Conclusion

Spectrum-based Fault Localization (SFL)

- **Spectra:**
  Cones of faulty and correct output variables

	A	B	C	D	E	F
1		**week 1**	**week 2**	**Total**	**$/h**	**Gross Pay**
2	Green	23	31	=SUM(B2)	15	=D2*E2
3	Jones	35	34	=SUM(B3:C3)	17	=D3*E3
4	**Total**	=SUM(B2:B3)	=SUM(C2:C3)	=SUM(D2:D3)		

CONE(F2) = {B2,D2,E2,F2}

CONE(D4) = {B2,D2,B3,C3,D3,D4}

CONE(B4) = {B2,B3,B4}

CONE(C4) = {C2,C3,C4}

CONE(F3) = {B3,C3,D3,E3,F3}

Hofer, Riboira, Wotawa, Abreu, Getzner: „On the Empirical Evaluation of Fault Localization Techniques for Spreadsheets“
16th Int. Conference on Fundamental Approaches to Software Engineering (FASE), March 2013, Rome, Italy

Spectrum-based Fault Localization (SFL)

- **Spectra:**
  Cones of faulty and correc

	A	B	C
1		**week 1**	**week 2**
2	Green	23	31
3	Jones	35	34
4	**Total**	=SUM(B2:B3)	=SUM(C2:C

CONE(F2) = {B2,D2,E2,F2}

CONE(D4) = {B2,D2,B3,C3,D3,D4}

CONE(B4) = {B2,B3,B4}

CONE(C4) = {C2,C3,C4}

CONE(F3) = {B3,C3,D3,E3,F3}

	F2	D4	B4	C4	F3	Coef.	Rank.
B2	●	●	●			0.816	2
B3		●	●		●	0.408	7
B4			●			-	
C2				●		-	
C3		●		●	●	0.408	7
C4				●		-	
D2	●	●				1.000	1
D3		●			●	0.500	6
D4		●				0.707	3
E2	●					0.707	3
E3					●	-	
F2	●					0.707	3
F3					●	-	
Error	●	●					

Hofer, Riboira, Wotawa, Abreu, Getzner: „On the Empirical Evaluation of Fault Localization Techniques for Spreadsheets“
16th Int. Conference on Fundamental Approaches to Software Engineering (FASE), March 2013, Rome, Italy

Outline

- Running Example & Definitions
- Debugging Methods
  - Spectrum-Based Fault Localization
  - Spectrum-Enhanced Dynamic Slicing
  - Constraint-Based Debugging
- Evaluation
- Future Work & Conclusion

Spectrum-enhanced Dynamic Slicing (SENDYS)

Slices/Cones as Conflict Sets → Hitting Sets → Minimal Diagnoses → Probabilities → Diagnoses Likelihood

Similarity Coefficients → Probabilities

CONE(F2)= {B2,D2,E2,F2}

CONE(D4)= {B2,D2,B3,C3,D3,D4}

Single fault diagnoses:
{B2}, {D2}

Double fault diagnoses:
{B3,E2}, {B3,F2}, {C3,E2}, {C3,F2}, {D3,E2}, {D3,F2}, {D4,E2}, {D4,F2}

p({B2}) = 0.231
p({D2}) = 0.289
p({B3,E2}= 0.009
…

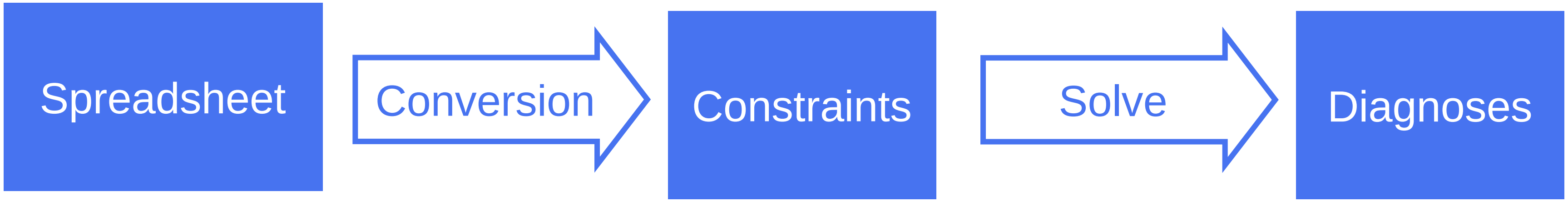
	F2	D4	B4	C4	F3	Coef.	Rank.
B2	•	•	•			0.816	2
B3		•	•		•	0.408	7
B4			•			-	
C2				•		-	
C3		•		•	•	0.408	7
C4				•		-	
D2	•	•				1.000	1
D3		•			•	0.500	6
D4		•				0.707	3
E2	•					0.707	3
E3					•	-	
F2	•					0.707	3
F3					•	-	
Error	•	•					

Hofer, Riboira, Wotawa, Abreu, Getzner: „On the Empirical Evaluation of Fault Localization Techniques for Spreadsheets“
16th Int. Conference on Fundamental Approaches to Software Engineering (FASE), March 2013, Rome, Italy

Outline

- Running Example & Definitions
- Debugging Methods
  - Spectrum-Based Fault Localization
  - Spectrum-Enhanced Dynamic Slicing
  - Constraint-Based Debugging
- Evaluation
- Future Work & Conclusion

Constraint-based Debugging (ConBug)



	D
1	**Total**
2	=SUM(B2)
3	=SUM(B3:C3)
4	=SUM(D2:D3)

AB(f_D2) ∨ (D2 == B2)
AB(f_D3) ∨ (D3 == B3+C3)
AB(f_D4) ∨ (D4 == D2+D4)
…

Test case:
D4 == 123
…

For single faults:
SUM(f_D2, f_D3, f_D4)==1

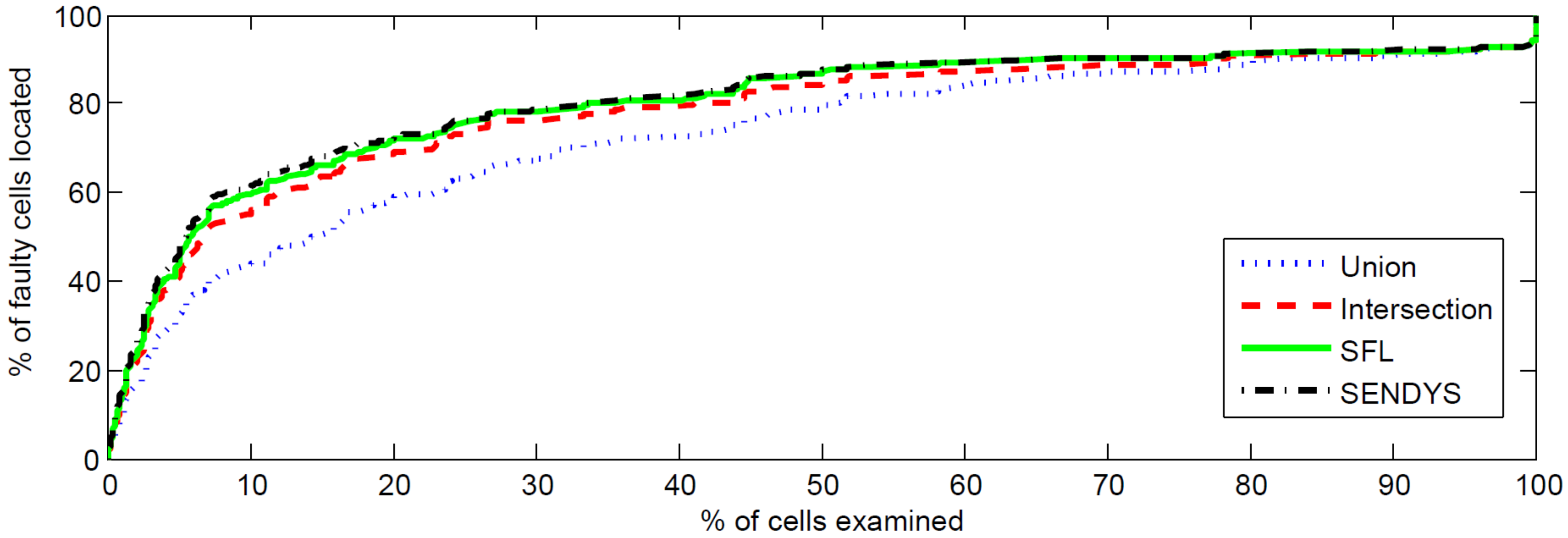
For complete example:

{D2}, {D4,F2}, …

Hofer, Riboira, Wotawa, Abreu, Getzner: „On the Empirical Evaluation of Fault Localization Techniques for Spreadsheets“
16th Int. Conference on Fundamental Approaches to Software Engineering (FASE), March 2013, Rome, Italy

Empirical evaluation – Part 1

- Spreadsheets form the EUSES spreadsheet corpus
- Filter: Excel 5.0, death references,
  no input values, ≤ 5 formulas
- 622 automatically created mutants
- 6 to 4170 formulas / spreadsheet (Avg: 225)

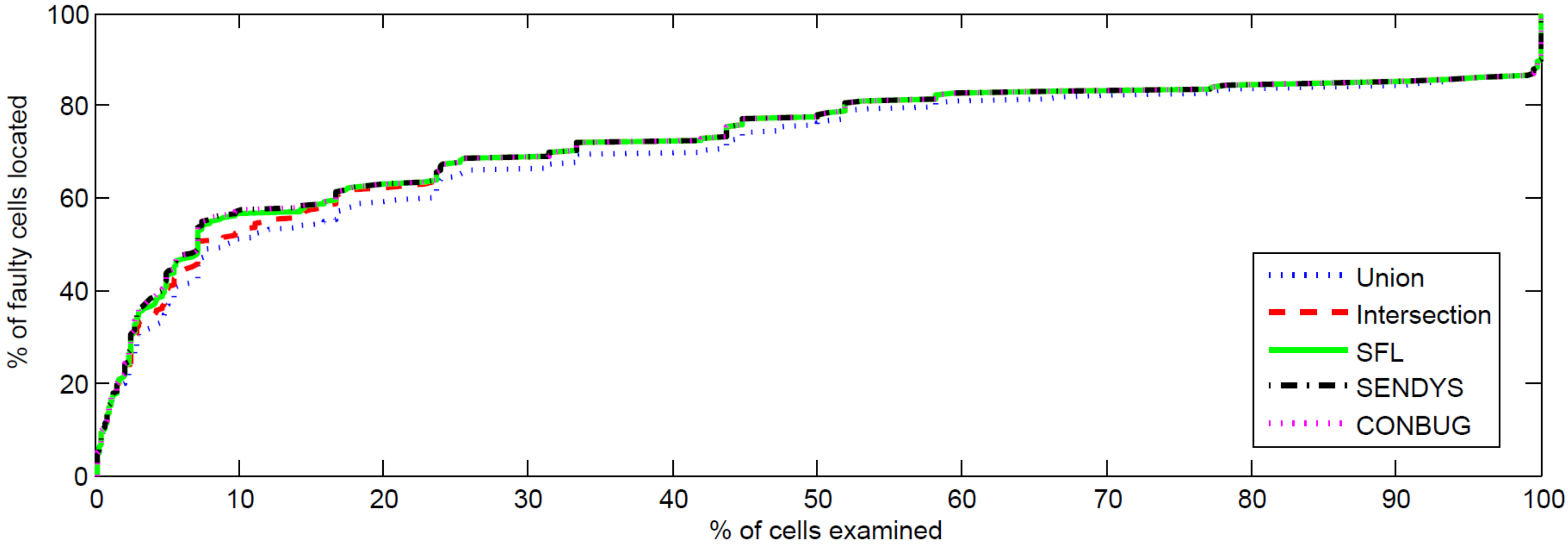


Hofer, Riboira, Wotawa, Abreu, Getzner: „On the Empirical Evaluation of Fault Localization Techniques for Spreadsheets“
16th Int. Conference on Fundamental Approaches to Software Engineering (FASE), March 2013, Rome, Italy

Empirical evaluation – Part 2

- Subset: 227 spreadsheets
- 6 to 2564 formulas / spreadsheet (Avg: 220)

Approach	Union	Intersec.	SFL	SENDYS	CONBUG
Avg. Time (ms)	14.0	13.9	15.0	63.9	631.7



Hofer, Riboira, Wotawa, Abreu, Getzner: „On the Empirical Evaluation of Fault Localization Techniques for Spreadsheets“
16th Int. Conference on Fundamental Approaches to Software Engineering (FASE), March 2013, Rome, Italy

Summary of the results

- SFL and SENDYS:
  - Outperform Intersection and Union
  - Performance dependent on the number of correct/incorrect output variables

- ConBug:
  - Significant computational overhead
  - Only small spreadsheets

Future work

- Improvements of ConBug
- Double faults
- Derive suggestions to use a specific method
- Provide solutions via mutations
- User acceptance study

Conclusion

Spectrum-based Fault Localization (SFL)

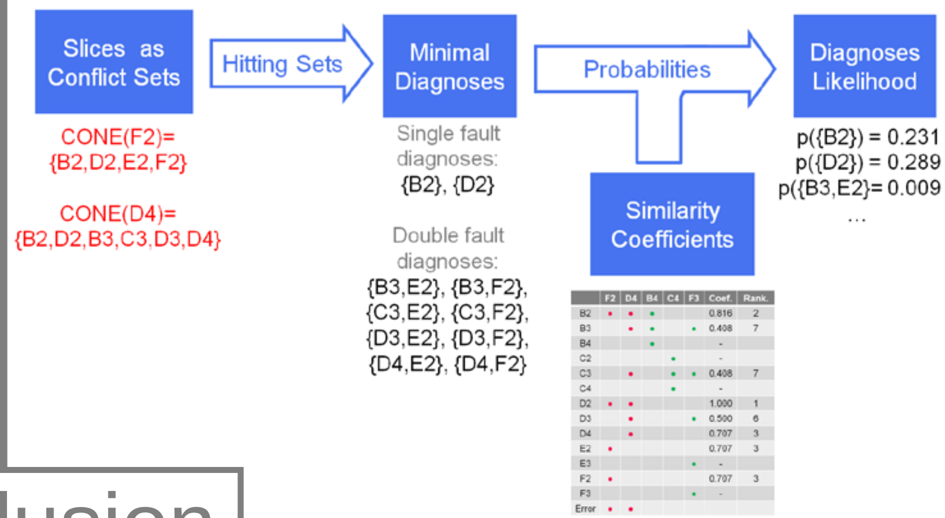
- **Spectra:**
  Cones of faulty and correct

	A	B	C
1		**week 1**	**week 2**
2	Green	23	31
3	Jones	35	34
4	**Total**	=SUM(B2:B3)	=SUM(C2:C

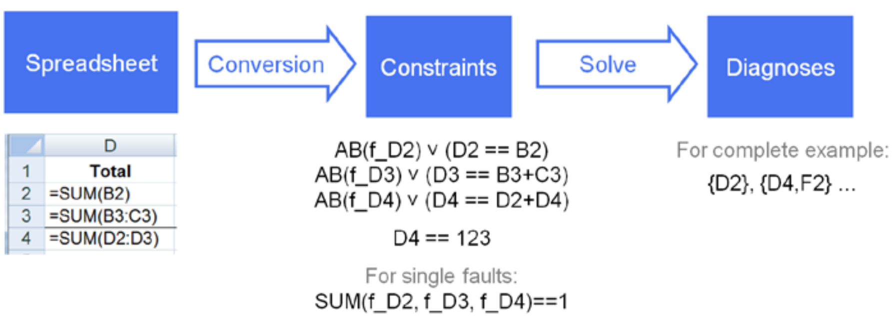
CONE(F2) = {B2,D2,E2,F2}
CONE(D4) = {B2,D2,B3,C3,D3,D4}
CONE(B4) = {B2,B3,B4}
CONE(C4) = {C2,C3,C4}
CONE(F3) = {B3,C3,D3,E3,F3}

	F2	D4	B4	C4	F3	Coef.	Rank.
B2	•	•	•			0.816	2
B3		•	•		•	0.408	7
B4			•			-	
C2				•		-	
C3		•		•	•	0.408	7
C4				•		-	
D2	•	•				1.000	1
D3		•			•	0.500	6
D4		•				0.707	3
E2	•					0.707	3
E3					•	-	
F2	•					0.707	3
F3					•	-	
Error	•	•					

Spectrum-enhanced Dynamic Slicing (SENDYS)

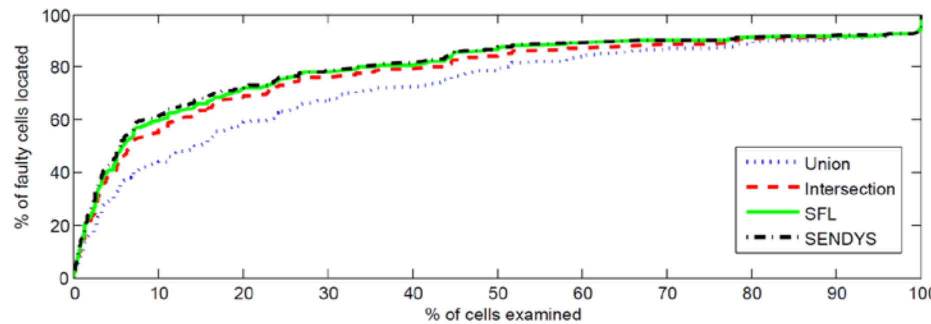


Constraint-based Debugging (ConBug)



Empirical evaluation – Part 1

- Spreadsheets form the EUSES spreadsheet corpus
- Filter: Excel 5.0 spreadsheets, ≤ 5 formulas
- 622 automatically created mutants
- 6 to 4170 formulas / spreadsheet (Avg: 225)



Hofer, Riboira, Wotawa, Abreu, Getzner: „On the Empirical Evaluation of Fault Localization Techniques for Spreadsheets“
16th Int. Conference on Fundamental Approaches to Software Engineering (FASE), March 2013, Rome, Italy

References

- Birgit Hofer and Franz Wotawa: Spectrum enhanced dynamic slicing for better fault localization. In The 20th European Conference on Artificial Intelligence (ECAI 2012).

- Birgit Hofer and Franz Wotawa. “Combining Slicing and Constraint Solving for Better Debugging: The CONBAS Approach.” In: Advances in Software Engineering, vol. 2012, Article ID 628571, 18 pages, 2012.

- Rui Abreu, André Riboira and Franz Wotawa: “Constraint-based Debugging of Spreadsheets”, Proceedings of the 15th Ibero-American Conference on Software Engineering, 2012.